### 第1章 ELKStack

　　对于日志来说，最常见的需求就是**收集、存储、查询、展示**，开源社区正好有相对应的开源项目：logstash（收集）、elasticsearch（存储+搜索）、kibana（展示），我们将这三个组合起来的技术称之为ELKStack，所以说ELKStack指的是Elasticsearch、Logstash、Kibana技术栈的结合。

### 1.1 Elasticsearch介绍

　　ElasticSearch是一个基于Lucene的搜索服务器。它提供了一个分布式多用户能力的全文搜索引擎，基于RESTful web接口。Elasticsearch是用Java开发的，并作为Apache许可条款下的开放源码发布，是第二流行的企业搜索引擎。设计用于云计算中，能够达到实时搜索，稳定，可靠，快速，安装使用方便。

### 1.1.1 Elasticsearch几个重要术语

**NRT**

elasticsearch是一个近似实时的搜索平台，从索引文档到可搜索有些延迟，通常为1秒。

**集群**

集群就是一个或**多个节点存储数据**，其中一个节点为主节点，这个主节点是可以通过选举产生的，并提供跨节点的联合索引和搜索的功能。集群有一个唯一性标示的名字，默认是elasticsearch，集群名字很重要，每个节点是基于集群名字加入到其集群中的。因此，确保在不同环境中使用不同的集群名字。一个集群可以只有一个节点。强烈建议在配置elasticsearch时，配置成集群模式。

**节点**

**节点就是一台单一的服务器**，是集群的一部分，存储数据并参与集群的索引和搜索功能。像集群一样，节点也是通过名字来标识，默认是在节点启动时随机分配的字符名。当然啦，你可以自己定义。该名字也蛮重要的，在集群中用于识别服务器对应的节点。

节点可以通过指定集群名字来加入到集群中。默认情况下，每个节点被设置成加入到elasticsearch集群。如果启动了多个节点，假设能自动发现对方，他们将会自动组建一个名为elasticsearch的集群。

**索引**

索引是有几分相似属性的一系列文档的集合。如nginx日志索引、syslog索引等等。索引是由名字标识，名字必须全部小写。这个名字用来进行索引、搜索、更新和删除文档的操作。

索引相对于关系型数据库的库。

**类型**

在一个索引中，可以定义一个或多个类型。类型是一个逻辑类别还是分区完全取决于你。通常情况下，一个类型被定于成具有一组共同字段的文档。如ttlsa运维生成时间所有的数据存入在一个单一的名为logstash-ttlsa的索引中，同时，定义了用户数据类型，帖子数据类型和评论类型。

类型相对于关系型数据库的表。

**文档**

文档是信息的基本单元，可以被索引的。文档是以JSON格式表现的。

在类型中，可以根据需求存储多个文档。

虽然一个文档在物理上位于一个索引，实际上一个文档必须在一个索引内被索引和分配一个类型。

文档相对于关系型数据库的列。

**分片和副本**

　　在实际情况下，索引存储的数据可能超过单个节点的硬件限制。如一个十亿文档需1TB空间可能不适合存储在单个节点的磁盘上，或者从单个节点搜索请求太慢了。为了解决这个问题，elasticsearch提供将索引分成多个分片的功能。当在创建索引时，可以定义想要分片的数量。每一个分片就是一个全功能的独立的索引，可以位于集群中任何节点上。

分片的两个最主要原因：

a、水平分割扩展，增大存储量

b、分布式并行跨分片操作，提高性能和吞吐量

分布式分片的机制和搜索请求的文档如何汇总完全是有elasticsearch控制的，这些对用户而言是透明的。

网络问题等等其它问题可以在任何时候不期而至，为了健壮性，强烈建议要有一个故障切换机制，无论何种故障以防止分片或者节点不可用。为此，elasticsearch让我们将索引分片复制一份或多份，称之为分片副本或副本。

**副本也有两个最主要原因：**

* 高可用性，以应对分片或者节点故障。出于这个原因，分片副本要在不同的节点上。
* 提供性能，增大吞吐量，搜索可以并行在所有副本上执行。

　　总之，每一个索引可以被分成多个分片。索引也可以有0个或多个副本。复制后，每个索引都有主分片(母分片)和复制分片(复制于母分片)。分片和副本数量可以在每个索引被创建时定义。索引创建后，可以在任何时候动态的更改副本数量，但是，不能改变分片数。

　　默认情况下，elasticsearch为每个索引分片5个主分片和1个副本，这就意味着集群至少需要2个节点。索引将会有5个主分片和5个副本(1个完整副本)，每个索引总共有10个分片。

　　每个elasticsearch分片是一个Lucene索引。一个单个Lucene索引有最大的文档数LUCENE-5843, 文档数限制为2147483519(MAX\_VALUE – 128)。 可通过\_cat/shards来监控分片大小。

### 1.2 logstash

### 1.2.1 LogStash 介绍

LogStash由JRuby语言编写，基于消息（message-based）的简单架构，并运行在Java虚拟机（JVM）上。不同于分离的代理端（agent）或主机端（server），LogStash可配置单一的代理端（agent）与其它开源软件结合，以实现不同的功能。

### 1.2.2 LogStash的四大组件

* Shipper：发送事件（events）至LogStash；通常，远程代理端（agent）只需要运行这个组件即可；
* Broker and Indexer：接收并索引化事件；
* Search and Storage：允许对事件进行搜索和存储；
* Web Interface：基于Web的展示界面

　　正是由于以上组件在**LogStash架构中可独立部署，才提供了更好的集群扩展性。**

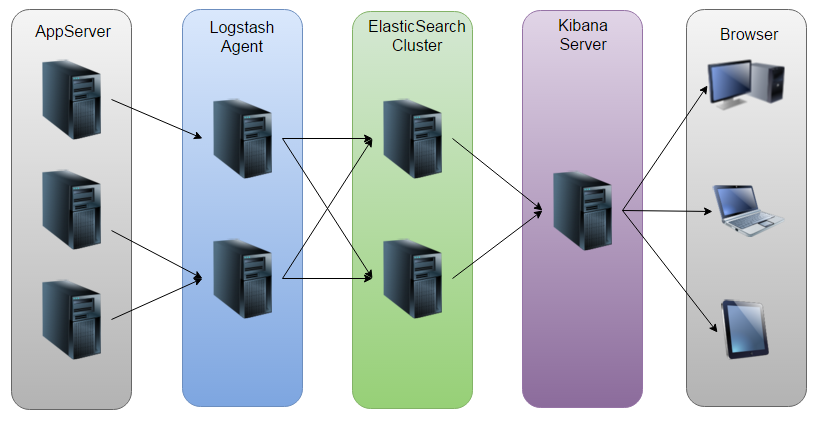
### 1.2.3 LogStash主机分类

* 代理主机（agent host）：作为事件的传递者（shipper），将各种日志数据发送至中心主机；只需运行Logstash 代理（agent）程序；
* 中心主机（central host）：可运行包括中间转发器（Broker）、索引器（Indexer）、搜索和存储器（Search and Storage）、Web界面端（Web  Interface）在内的各个组件，以实现对日志数据的接收、处理和存储。

## 1.3 Kibana

Logstash是一个完全开源的工具，他可以对你的日志**进行收集、分析**，并将其存储供以后使用（如，搜索），您可以使用它。说到搜索，logstash带有一个web界面，**搜索和展示所有日志。**

### 1.4 工作原理图



如图：Logstash收集AppServer产生的Log，并存放到ElasticSearch集群中，而Kibana则从ES集群中查询数据生成图表，再返回给Browser。

### 1.5 使用ELK必要性（解决运维痛点）

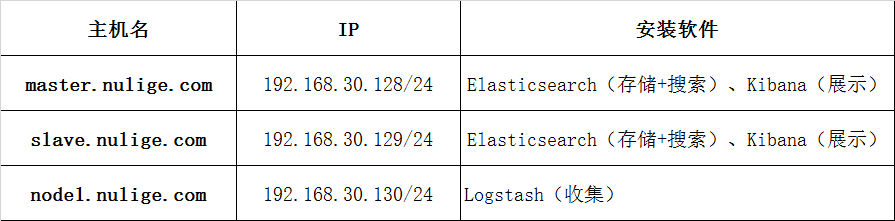
* 开发人员不能登录线上服务器查看详细日志
* 各个系统都有日志，日志数据分散难以查找
* 日志数据量大，查询速度慢，或者数据不够实时

### 第2章 ELK部署

#### 2.1 环境准备

[root@server elasticsearch]# cat /etc/redhat-release   
CentOS release 6.6 (Final)  
[root@server elasticsearch]# uname -r  
2.6.32-504.el6.x86\_64

配置表：



备注：另外可以增添服务器组，集群。

#### 服务器配置表如下：



## 2.2 Elasticsearch

Elasticsearch需要Java环境，可以直接使用yum安装。也可以从Oracle官网下载JDK进行安装。开始之前要确保JDK正常安装并且环境变量也配置正确。

### 2.2.1 安装JDK

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7 | [root@m01 ~]# yum install -y java    [root@m01 ~]# java -version    openjdk version "1.8.0\_121"  OpenJDK Runtime Environment (build 1.8.0\_121-b13)  OpenJDK 64-Bit Server VM (build 25.121-b13, mixed mode) |

### 2.2.2 添加elasticsearch、logstash、kibana的yum仓库

1、下载并安装GPG key

|  |  |
| --- | --- |
| 1 | rpm --import https://packages.elastic.co/GPG-KEY-elasticsearch |

2、添加elasticsearch.repo源 (或者直接从官网下载rpm包上传到服务器上面安装)

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15 | cat >/etc/yum.repos.d/elasticsearch.repo<<EOF    [elasticsearch-2.x]    name=Elasticsearch repository for 2.x packages    baseurl=http://packages.elastic.co/elasticsearch/2.x/centos    gpgcheck=1    gpgkey=http://packages.elastic.co/GPG-KEY-elasticsearch    enabled=1    EOF |

3、安装elasticsearch （**存储+搜索）**

|  |  |
| --- | --- |
| 1 | [root@server elasticsearch]# yum install -y elasticsearch |

4、  修改 limits.conf

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9 | #在结尾添加如下两行代码  [root@server elasticsearch]#vi /etc/security/limits.conf  elasticsearch soft memlock unlimited  elasticsearch hard memlock unlimited    #检查  [root@server elasticsearch]# tail -2 /etc/security/limits.conf  elasticsearch soft memlock unlimited  elasticsearch hard memlock unlimited |

5、创建目录并授权

|  |  |
| --- | --- |
| 1  2  3 | [root@m01 ~]# mkdir -p /data/es-data    [root@m01 ~]# chown -R elasticsearch.elasticsearch /data/es-data/ |

6、配置Elasticsearch

[复制代码](javascript:void(0);)

cat >/etc/elasticsearch/elasticsearch.yml<<EOF

cluster.name: elk-cluter  
node.name: linux-node1  
path.data: /data/es-data  
path.logs: /var/log/elasticsearch/  
bootstrap.mlockall: true  
network.host: 192.168.30.129  
http.port: 9200

EOF

[复制代码](javascript:void(0);)

#检查配置文件

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15 | [root@m01 ~]# grep '^[a-z]' /etc/elasticsearch/elasticsearch.yml    cluster.name: elk-cluter            # 集群的名称，判别节点是否是统计集群    node.name: linux-node1              # 节点的名称    path.data: /data/es-data            # 数据存放路径    path.logs: /var/log/elasticsearch/  #  日志路径    bootstrap.mlockall: true            # 锁住内存，使内存不会在swap中使用    network.host:192.168.30.129         # 允许访问的IP或域名    http.port: 9200                     # 端口 |

7、启动Elasticsearch

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8 | [root@server elasticsearch]# /etc/init.d/elasticsearch start  正在启动 elasticsearch：OpenJDK 64-Bit Server VM warning: If the number of processors is expected to increase from one, then you should configure the number of parallel GC threads appropriately using -XX:ParallelGCThreads=N  [确定]    #检查启动  [root@server elasticsearch]# ps -ef|grep elasticsearch  498        1422      1 96 21:50 ?        00:00:15 /usr/bin/java -Xms256m -Xmx1g -Djava.awt.headless=true -XX:+UseParNewGC -XX:+UseConcMarkSweepGC -XX:CMSInitiatingOccupancyFraction=75 -XX:+UseCMSInitiatingOccupancyOnly -XX:+HeapDumpOnOutOfMemoryError -XX:+DisableExplicitGC -Dfile.encoding=UTF-8 -Djna.nosys=true -Des.path.home=/usr/share/elasticsearch -cp /usr/share/elasticsearch/lib/elasticsearch-2.4.4.jar:/usr/share/elasticsearch/lib/\* org.elasticsearch.bootstrap.Elasticsearch start -p /var/run/elasticsearch/elasticsearch.pid -d -Des.default.path.home=/usr/share/elasticsearch -Des.default.path.logs=/var/log/elasticsearch -Des.default.path.data=/var/lib/elasticsearch -Des.default.path.conf=/etc/elasticsearch  root       1440   1046  0 21:51 pts/0    00:00:00 grep elasticsearch |

8、测试网站看是否能访问

备注：如果是用域名，请先做域名解析，否则无法访问成功，并要加端口。（不要用内网IP，无法访问成功，除非你用VPN）

|  |  |
| --- | --- |
| 1  2 | #windows系统，在hosts文件里，配置域名解析  192.168.30.129:9200 server.nulige.com<br>192.168.30.128:9200 slave.nulige.com |

在linux中访问测试：

|  |  |
| --- | --- |
| 1 | #可以用ip访问 |

[root@server ~]# curl 192.168.30.129:9200  
{  
"name" : "linux-node1",  
"cluster\_name" : "elk-cluter",  
"cluster\_uuid" : "lBOcpanYR9WQnM3zZQ7org",  
"version" : {  
"number" : "2.4.4",  
"build\_hash" : "fcbb46dfd45562a9cf00c604b30849a6dec6b017",  
"build\_timestamp" : "2017-01-03T11:33:16Z",  
"build\_snapshot" : false,  
"lucene\_version" : "5.5.2"  
},  
"tagline" : "You Know, for Search"  
}

在windows系统中访问



### 2.2.3 Elasticsearch插件

1、安装Elasticsearch集群管理插件

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7 | [root@server elasticsearch]# /usr/share/elasticsearch/bin/plugin install mobz/elasticsearch-head  -> Installing mobz/elasticsearch-head...  Trying https://github.com/mobz/elasticsearch-head/archive/master.zip ...  Downloading ...................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................DONE  Verifying https://github.com/mobz/elasticsearch-head/archive/master.zip checksums if available ...  NOTE: Unable to verify checksum for downloaded plugin (unable to find .sha1 or .md5 file to verify)  Installed head into /usr/share/elasticsearch/plugins/head   #安装路径 |

在线安装（若无法下载，请使用离线安装方法）：

**离线安装：**

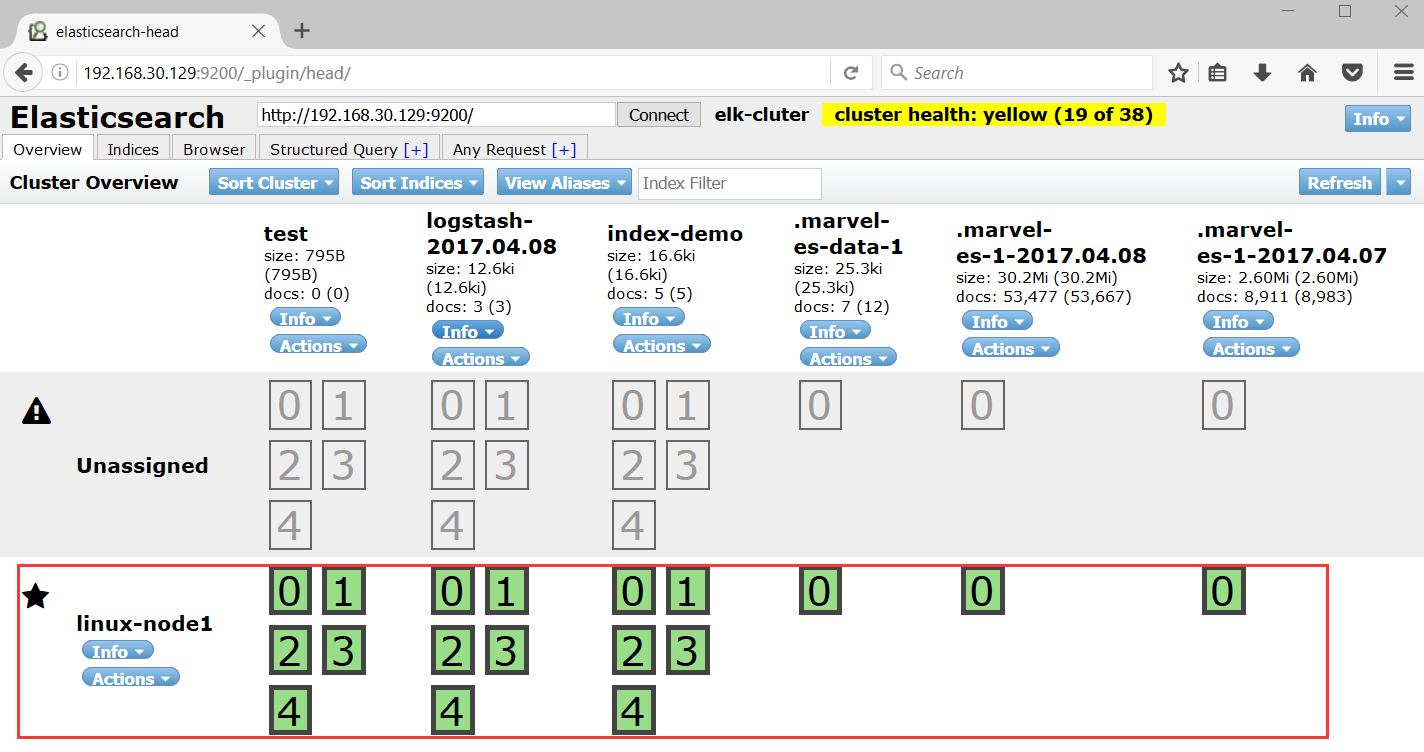
|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14 | [root@server elasticsearch]# wget https://github.com/mobz/elasticsearch-head/archive/master.zip    [root@server elasticsearch]# /usr/share/elasticsearch/bin/plugin install file:/root/master.zip  -> Installing from file:/root/master.zip...    Trying file:/root/master.zip ...    Downloading .........DONE    Verifying file:/root/master.zip checksums if available ...    NOTE: Unable to verify checksum for downloaded plugin (unable to find .sha1 or .md5 file to verify)    Installed head into /usr/share/elasticsearch/plugins/head |

插件安装帮助信息：

[+ View Code](https://www.cnblogs.com/nulige/p/6680336.html)

### 2.2.4 访问head集群插件

|  |  |
| --- | --- |
| 1 | http://192.168.30.129:9200/\_plugin/head/ |

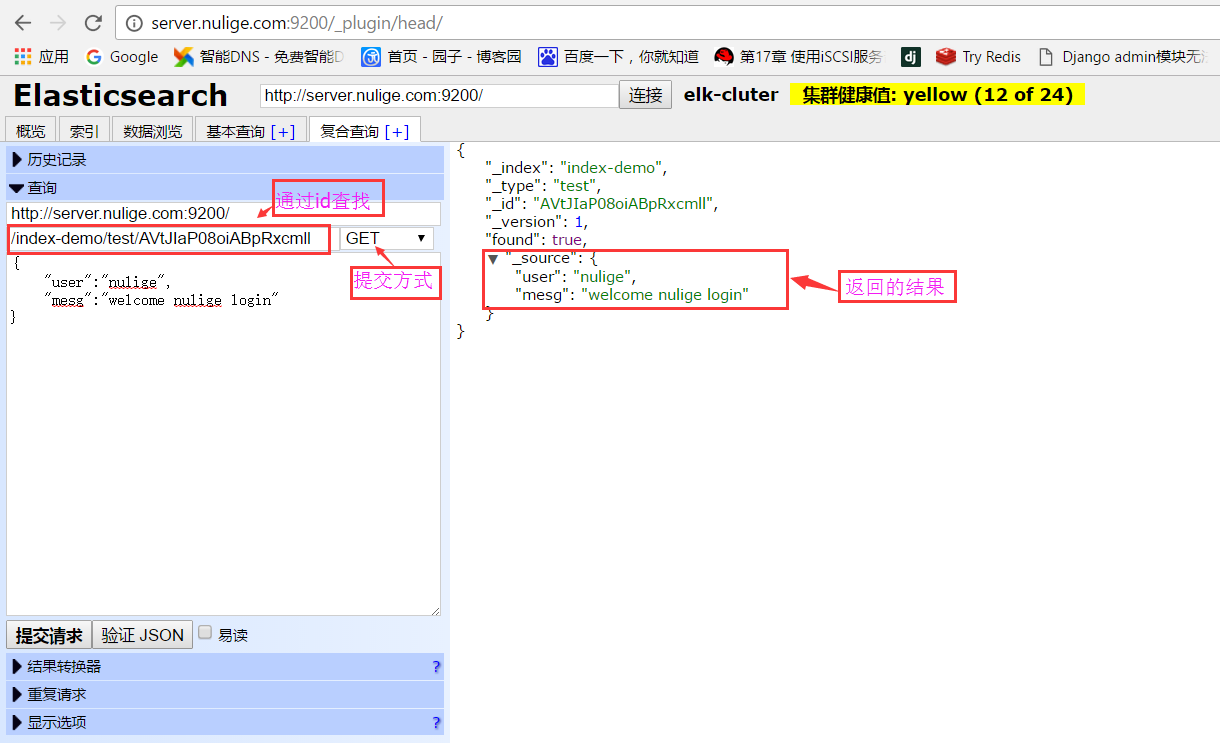


示例1：用复合查询

通过post提交，设置查询条件，得到结果



通过get的方式，用id的值，查询出结果：



示例2：基本查询

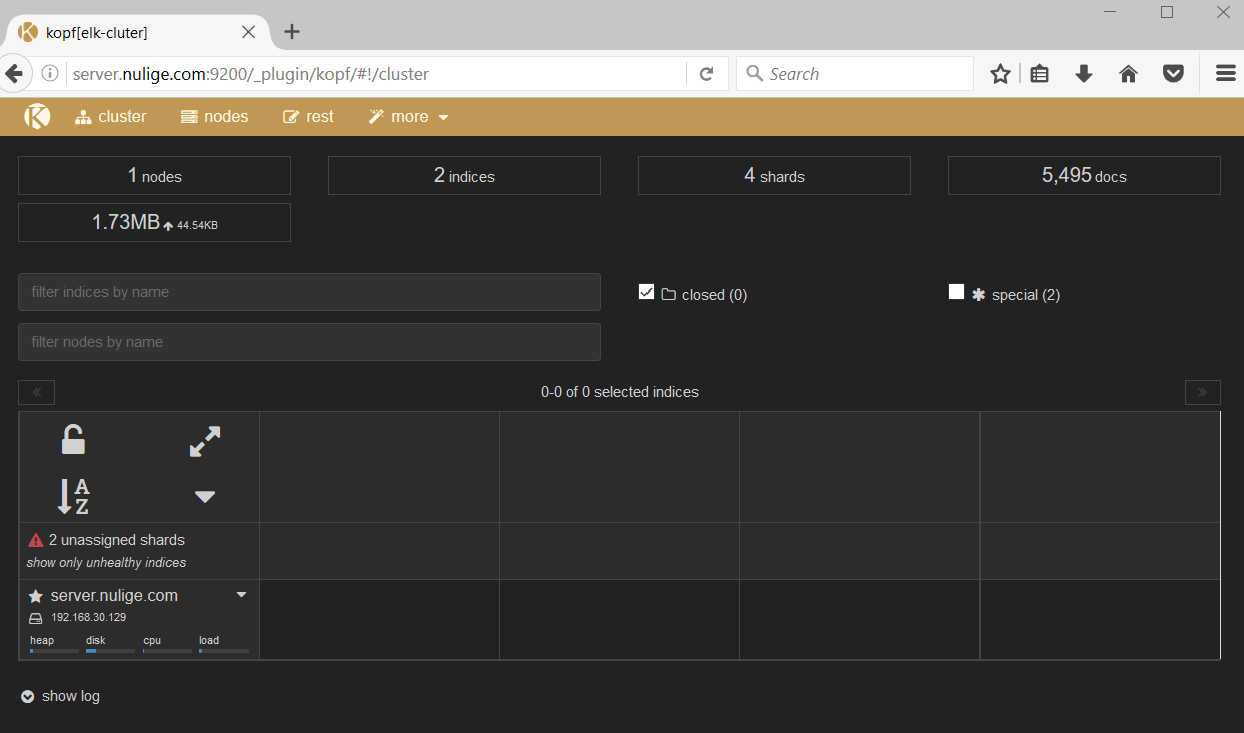


### 2.2.5 ****server.nulige.com（主）****安装监控插件（根据需求选装）

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7 | [root@server head]# /usr/share/elasticsearch/bin/plugin install lmenezes/elasticsearch-kopf  -> Installing lmenezes/elasticsearch-kopf...  Trying https://github.com/lmenezes/elasticsearch-kopf/archive/master.zip ...  Downloading ..................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................DONE  Verifying https://github.com/lmenezes/elasticsearch-kopf/archive/master.zip checksums if available ...  NOTE: Unable to verify checksum for downloaded plugin (unable to find .sha1 or .md5 file to verify)  Installed kopf into /usr/share/elasticsearch/plugins/kopf |

访问网站地址：

http://server.nulige.com:9200/\_plugin/kopf/#!/cluster



### 2.2.6 elasticsearch集群

如果要配置集群，另外配置集群名相同的节点，通过组播进行通信，会通过cluster进行查找，如果无法通过组播查询，修改成单播即可。

配置文件参考如下：

grep "^[a-Z]" /etc/elasticsearch/elasticsearch.yml

内容：

cluster.name: elk-cluter  
node.name: linux-node2  
path.data: /data/es-data  
path.logs: /var/log/elasticsearch/  
bootstrap.mlockall: true  
network.host: 192.168.30.128  
http.port: 9200  
discovery.zen.ping.unicast.hosts: ["192.168.30.129","192.168.30.128"]  #单播（配置一台即可，生产可以使用组播方式）

 修改方法：

#停止服务

[root@slave ~]# /etc/init.d/elasticsearch stop

#启动服务

[root@slave ~]# /etc/init.d/elasticsearch start

备注：先停止服务，再修改配置文件，再启动服务，可在集群管理插件页面查到相关信息。

　　在浏览器中查看分片信息，一个索引默认被分成了5个分片，每份数据被分成了五个分片（可以调节分片数量），下图中外围带绿色框的为主分片，不带框的为副本分片，主分片丢失，副本分片会复制一份成为主分片，起到了高可用的作用，主副分片也可以使用负载均衡加快查询速度，但是如果主副本分片都丢失，则索引就是彻底丢失。

## 2.3 LogStash部署与配置

先安装JDK，也可以下载安装Oracle的JDK或者使用 yum安装openjdk。

|  |  |
| --- | --- |
| 1 | yum install -y java |

### 2.3.1 安装logstash （作用：收集日志）

1、下载并安装GPG key

|  |  |
| --- | --- |
| 1 | rpm --import https://packages.elastic.co/GPG-KEY-elasticsearch |

2、添加yum仓库

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15 | cat >/etc/yum.repos.d/logstash.repo<<EOF    [logstash-2.3]    name=Logstash repository for 2.3.x packages    baseurl=https://packages.elastic.co/logstash/2.3/centos    gpgcheck=1    gpgkey=https://packages.elastic.co/GPG-KEY-elasticsearch    enabled=1    EOF |

3、安装logstash

|  |  |
| --- | --- |
| 1 | yum install -y logstash |

### 2.3.2 logstash入门

启动一个logstash，-e：在命令行执行；input输入，stdin标准输入，是一个插件；output输出，stdout：标准输出

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14 | [root@node1 tools]# /opt/logstash/bin/logstash -e 'input { stdin {} } output { stdout{codec => rubydebug} }'  OpenJDK 64-Bit Server VM warning: If the number of processors is expected to increase from one, then you should configure the number of parallel GC threads appropriately using -XX:ParallelGCThreads=N  hello   #<===会停在这里，等待让你输入内容。    #等待一会，才会返回下面结果：    Settings: Default pipeline workers: 1  Pipeline main started  {         "message" => "hello",        "@version" => "1",      "@timestamp" => "2017-04-08T04:44:14.153Z",            "host" => "node1.nulige.com"  } |

#他会显示出内容，版本号，时间，主机名

使用rubydebug显示详细输出，codec为一种编码器

使用logstash讲信息写入到Elasticsearch：

|  |  |
| --- | --- |
| 1  2  3  4  5 | [root@node1 tools]# /opt/logstash/bin/logstash -e 'input { stdin{} } output { elasticsearch { hosts => ["server.nulige.com:9200"] } }'  OpenJDK 64-Bit Server VM warning: If the number of processors is expected to increase from one, then you should configure the number of parallel GC threads appropriately using -XX:ParallelGCThreads=N  hello  #输入一条内容  Settings: Default pipeline workers: 1  Pipeline main started |

每一条输出的内容称为一个事件，多个相同的输出的内容合并到一起称为：一个事件

### 2.3.3 收集nginx访问日志

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27 | #修改nginx日志格式为json        log\_format json '{ "@timestamp": "$time\_local", '             '"@fields": { '             '"remote\_addr": "$remote\_addr", '             '"remote\_user": "$remote\_user", '             '"body\_bytes\_sent": "$body\_bytes\_sent", '             '"request\_time": "$request\_time", '             '"status": "$status", '             '"request": "$request", '             '"request\_method": "$request\_method", '             '"http\_referrer": "$http\_referer", '             '"body\_bytes\_sent":"$body\_bytes\_sent", '             '"http\_x\_forwarded\_for": "$http\_x\_forwarded\_for", '             '"http\_user\_agent": "$http\_user\_agent" } }'; |

## 2.4 安装Kibana

1、添加kibana.repo源

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15 | cat >/etc/yum.repos.d/kibana.repo<<EOF    [kibana-4.5]    name=Kibana repository for 4.5.x packages    baseurl=http://packages.elastic.co/kibana/4.5/centos    gpgcheck=1    gpgkey=http://packages.elastic.co/GPG-KEY-elasticsearch    enabled=1    EOF |

2、安装kibana

|  |  |
| --- | --- |
| 1  2 | #展示效果  [root@server ~]# yum install -y kibana |

3、修改kibana的配置文件,开启下面选项

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | [root@server ~]# vi /opt/kibana/config/kibana.yml  server.port: 5601  server.host: "0.0.0.0"  elasticsearch.url: "http://server.nulige.com:9200"  elasticsearch.preserveHost: true  kibana.index: ".kibana" |

#检查

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | [root@server ~]# egrep -v "^$|#" /opt/kibana/config/kibana.yml  server.port: 5601  server.host: "0.0.0.0"  elasticsearch.url: "http://server.nulige.com:9200"  elasticsearch.preserveHost: true  kibana.index: ".kibana" |

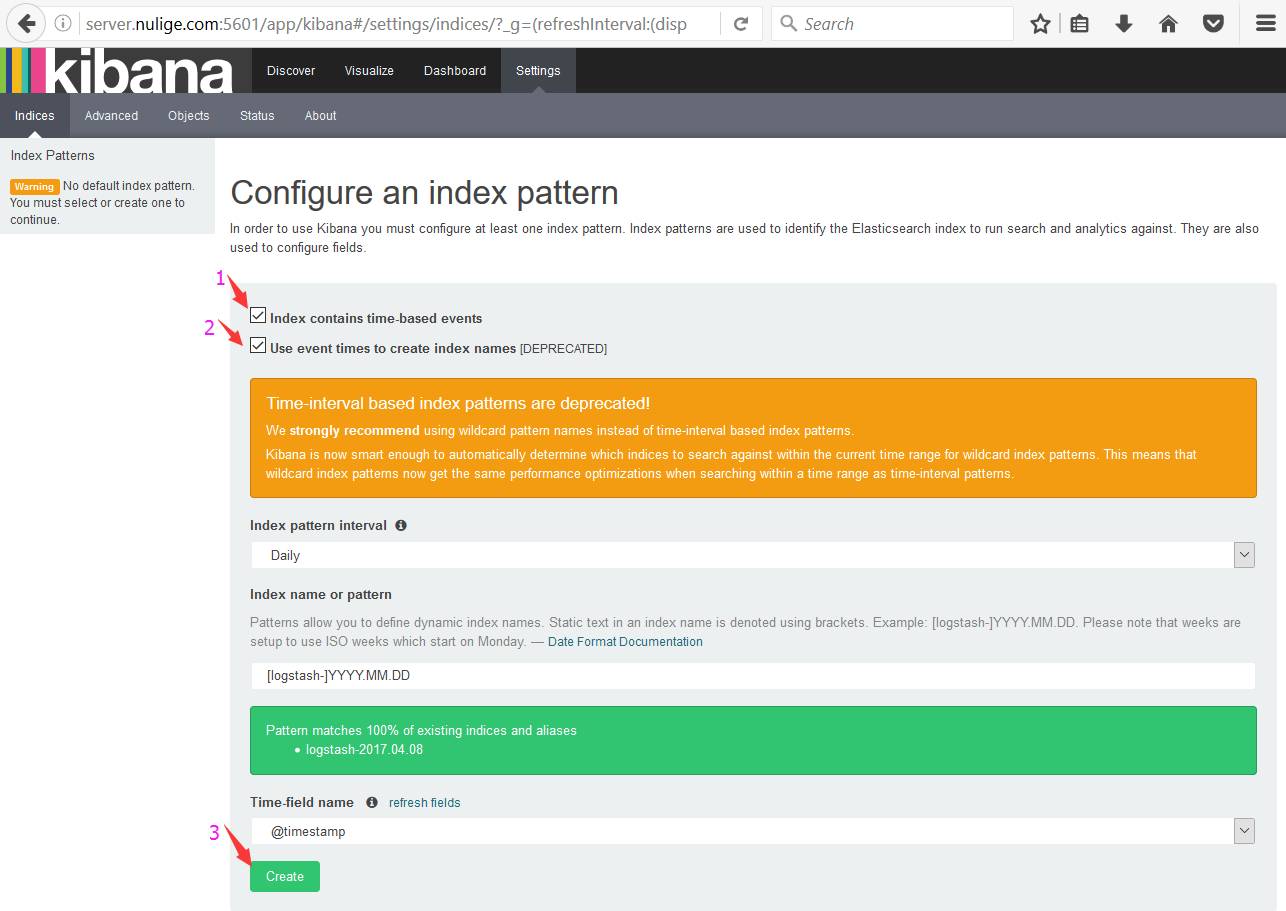
4、启动服务并检查

|  |  |
| --- | --- |
| 1  2  3  4  5 | [root@server ~]# /etc/init.d/kibana restart  kibana started<br>  [root@server ~]# ps -ef|grep kibana  kibana     3559      1 55 14:16 pts/0    00:00:07 /opt/kibana/bin/../node/bin/node /opt/kibana/bin/../src/cli  root       3571   2931  0 14:17 pts/0    00:00:00 grep kibana |

5、访问网站：

|  |  |
| --- | --- |
| 1 | http://server.nulige.com:5601 |

6、创建默认索引



7、展示今天的数据

